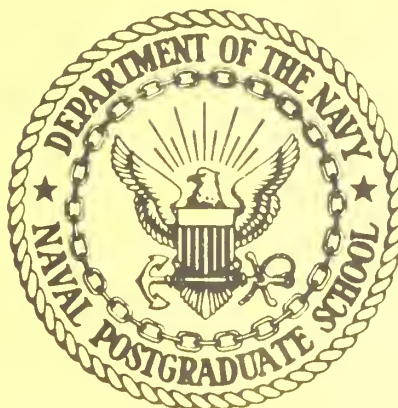NPS55-82-003

NAVAL POSTGRADUATE SCHOOL
Monterey, California

STATIONARY DISCRETE AUTOREGRESSIVE-MOVING

AVERAGE TIME SERIES GENERATED BY MIXTURES

by

P. A. Jacobs

P. A. W. Lewis

January 1982

Approved for public release; distribution unlimited

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral J. J. Ekelund   David A. Schrady
Superintendent   Acting Provost

Unclassified

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1 REPORT NUMBER NPS55-82-003	2 GOVT ACCESSION NO.	3 RECIPIENT'S CATALOG NUMBER
4 TITLE (and Subtitle) STATIONARY DISCRETE AUTOREGRESSIVE-MOVING AVERAGE TIME SERIES GENERATED BY MIXTURES		5. TYPE OF REPORT & PERIOD COVERED Technical
		6 PERFORMING ORG. REPORT NUMBER
7 AUTHOR(s) P. A. Jacobs P. A. W. Lewis		8 CONTRACT OR GRANT NUMBER(s) NSF-ENG-79-01438 NSF-ENG-79-10825
9 PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940		10 PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N;RR014-05-01 N0001481WR10001 61153N;RR014-05-01 N0001482AF00001
11 CONTROLLING OFFICE NAME AND ADDRESS Chief of Naval Research Arlington, VA 22217 National Science Foundation Washington, DC 20550		12 REPORT DATE January 1982
		13 NUMBER OF PAGES 34
14 MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15 SECURITY CLASS. (of this report) Unclassified
		15a DECLASSIFICATION/DOWNGRADING SCHEDULE

16 DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17 DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18 SUPPLEMENTARY NOTES

19 KEY WORDS (Continue on reverse side if necessary and identify by block number)

stationary processes; DARMA(p,N+1); NDARMA(p,N); mover-stayer models; mixtures; discrete time series; Markov chains; DAR(1) process

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Two simple stationary processes of discrete random variables with arbitrarily chosen first-order marginal distributions, DARMA(p,N+1) and NDARMA(p,N), are given. The correlation structure of these processes mimics that of the usual linear ARMA(p,q) processes. The relationship of these processes to mover-stayer models, and to models for discrete time series given separately by Lindqvist and Pegram is discussed. Ad-hoc nonparametric estimators for the parameters in the DARMA(p,N+1) and NDARMA(p,N) are given. A simulation study

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
: JAN 73
S/N 0102-014-6601

Unclassified

shows them to be as good as maximum likelihood estimators for the first-order autoregressive case, and to be much simpler to compute than the maximum likelihood estimators.

STATIONARY DISCRETE AUTOREGRESSIVE-MOVING AVERAGE

TIME SERIES GENERATED BY MIXTURES

by

P. A. Jacobs

and

P. A. W. Lewis

Operations Research Department
Naval Postgraduate School
Monterey, California 93940

ABSTRACT

Two simple stationary processes of discrete random variables with arbitrarily chosen first-order marginal distributions, DARMA(p,N+l) and NDARMA(p,N), are given. The correlation structure of these processes mimics that of the usual linear ARMA(p,q) processes. The relationship of these processes to mover-stayer models, and to models for discrete time series given separately by Lindqvist and Pegram is discussed. Ad-hoc nonparametric estimators for the parameters in the DARMA(p,N+l) and NDARMA(p,N) are given. A simulation study shows them to be as good as maximum likelihood estimators for the first-order autoregressive case, and to be much simpler to compute than the maximum likelihood estimators.

1.   INTRODUCTION

Discrete time series arise in many different contexts. For example the exact arrival times in an arrival process are usually not measured. Instead the number of arrivals in successive time intervals are given. This is the case with the statistics published by the Center for Disease Control on the incidence of various diseases in the United States. The data are given as the number of occurrences of each disease in successive days. If the time intervals are short enough and the arrival process is orderly, then the resulting time series is approximately binary. In other instances the process that is being measured is continuous but the data is quantized in recording. For example, the amount of rainfall in a day (24 hours) at a location, given that some occurs, is a continuous random variable; however, it is often recorded to the nearest one-hundredth or one-tenth of an inch. Also, since a rainfall series will often contain many zeros (no rain), an analysis is often made of successive wet and dry days which is a binary time series [cf. Buishand, 1978]. An economic imperative for modelling and predicting the binary rainfall series is that it is the primary concomittant variable for predicting volume of business done in some establishments on successive days.

Markov chains have been used as models for stationary discrete time series. However, they are overparametrized for statistical purposes. Further, the data to be modelled can often be shown to be non-Markovian, or at least not first-order Markovian. Higher order Markov chains can be used but this only aggravates the problem of overparametrization.

In the past several years various parametrically simple models have been proposed for stationary discrete time series. The models have as parameters the fixed, first-order marginal distribution of the time series and the correlation structure. In Jacobs and Lewis [1978a, 1978b] a simple scheme is given for obtaining a stationary sequence of discrete random variables with a given marginal probability mass function π and an autocorrelation structure like that of a mixed first-order autoregressive-(N+1)st-order moving average process. This DARMA(1, N+1) process has nonnegative correlations and a possibly countably infinite state space. The correlation structure is determined by parameters that are independent of the marginal distribution.

A special case of the DARMA(1, N+1) process with marginal probability mass function π is the DAR(1) process. This is a Markov chain with discrete state space $\mathbb{E}$ and with transition matrix

$$(1.1) \qquad\qquad P = \rho\underline{I} + (1-\rho)\underline{Q} \ ,$$

where $\underline{Q}$ is a matrix with $Q_{ij} = \pi(j)$ for $i,j \in \mathbb{E}$; $\underline{I}$ is the identity matrix with (i,j) element I_{ij} and $0 \le \rho < 1$. The correlation structure of a real valued DAR(1) i.e. one for which $\mathbb{E}$ is a subset of the real line, is that of a first-order autoregressive process with kth-order serial correlation equal to ρ^k. There is no limitation on π; a common and useful assumption is that it be Poisson and therefore have an infinite state space. The DAR(1) model with a finite state space is a special case of the mover-stayer model [Bhat, 1972, p. 302-9].

Lindqvist [1978] studied a real valued finite state space Markov chain with a transition function of the form

(1.2) $$\bar{P}_{ij} = cI_{ij} + (1-c) Q_{ij} \, ,$$

where $Q_{ij} = \pi(j)$, for i, $j \in \mathbb{E}$ as before. Since the state space $\mathbb{E}$ is finite, the constant c can take on some negative values with the constraint that

(1.3) $$\max_{1 \leq i \leq r} [1 - (1-\pi(i))^{-1}] \leq c \leq 1 \, ,$$

where r is the number of elements in the state space.

In Jacobs and Lewis [1978c], the DAR(1) process was extended to obtain a sequence of discrete random variables with pth order Markov dependence and given marginal distribution. The DAR(p) process is defined as follows. Let $\{V_n\}$ be a sequence of independent identically distributed random variables with $P\{V_n = 1\} = 1 - P\{V_n = 0\} = 1 - \rho$, $0 \leq \rho < 1$; $\{A_n\}$ is a sequence of independent identically distributed random variables taking values $\{1,2,...p\}$ with $P\{A_n = i\} = \alpha_i$, $i = 1,2,...,p$; and $\{Y_n\}$ is a sequence of independent identically distributed random variables with discrete state space $\mathbb{E}$ and $P\{Y_n = i\} = \pi(i)$. Let

(1.4) $$Z_n = V_n Z_{n-A_n} + (1-V_n)Y_n \, .$$

The process $\{Z_n\}$ is called a DAR(p) process. Note that by direct argument from (1.4)

3

(1.5) $P\{Z_{n+1} = j | Z_n = i_1, \ldots, Z_{n-p+1} = i_p\}$

$$= (1-\rho) \pi(j) + \sum_{k=1}^{p} \rho\alpha_k \varepsilon_j(i_k) \ ,$$

where $\varepsilon_j(i) = 1$ if $i = j$ and $\varepsilon_j(i) = 0$ otherwise. There is no limitation on the marginal probability mass function π.

Pegram [1980] considers a real-valued <u>finite state space</u> model $\{Z'_n\}$ which is a generalization of the DAR(p) model in that its conditional probabilities are of the form

(1.6) $P\{Z'_{n+1} = j | Z'_n = i_1, \ldots, Z'_{n-p+1} = i_p\}$

$$= [1 - \sum_{k=1}^{p} \phi_k] \pi(j) + \sum_{k=1}^{p} \phi_k \varepsilon_j(i_k) \ ,$$

where $\{\phi_k; k=1,\ldots,p\}$ are (possibly) negative constants. Note that although some of the constants ϕ_k may be negative, the admissible values for $\{\phi_k\}$ depends on the marginal distribution π. It was shown in Jacobs and Lewis, [1978c] that $\text{Corr}(Z_n, Z_{n+k})$, $k=1,2,\ldots$ for the real valued DAR(p) process are nonnegative. Pegram's model allows some of the correlations to be negative. The amount of negative correlation, as in Lindqvist's model, depends on the marginal distribution π.

In this paper we will consider models for real-valued stationary discrete time series whose nonnegative correlation structure is that of a mixed pth-order autoregressive and qth-order moving average process. Thus we have a generalization of all

of the preceding models. In Section 2 we will give definitions
of two such models, DARMA(p,N+1) (discrete mixed autoregressive-
moving average process with orders p and N+1 respectively)
and NDARMA(p,N). We briefly describe some of their properties
and suggest an estimate for the correlations. In Section 3
we describe in detail a simulation experiment that was done
to study the behavior of various estimators for the first order
serial correlation coefficients ρ of the DAR(1) model for
small and moderate sample sizes. In Section 4 some extensions
of the DARMA models are briefly discussed including one which
can have negative correlations. Throughout the remainder of
the paper we will assume that the NDARMA and DARMA processes
are real valued. They can in fact be used to model categorical
time series, but then numerical measures such as correlations
are meaningless.

2. The DARMA(p,N+1) and NDARMA(p,N) Processes.

In what follows we let $\{Y_n\}$ be a sequence of independent identically distributed random variables taking values in a real-valued discrete state space $\mathbb{E}$ with $P\{Y_n = i\} = \pi(i)$, $i \in \mathbb{E}$. Let $\{U_n\}$ and $\{V_n\}$ be independent sequences of independent random variables taking the values 0 and 1 with

(2.1) $P\{U_n = 1\} = \beta$ and $P\{V_n = 1\} = \rho$

for fixed $0 \le \beta \le 1$ and $0 \le \rho < 1$. Let $\{D_n\}$ be a sequence of independent identically distributed random variables taking values $0,1,2,\ldots,N$ with $P\{D_k = n\} = \delta_n$, $n = 0,1,\ldots,N$, and $\{A_n\}$ be a sequence of independent identically distributed random variables taking values $1,2,\ldots,p$ with $P\{A_k = n\} = \alpha_n$, $n = 1,2,\ldots,p$.

2.1 The DARMA(p,N+1) process.

The DARMA(p,N+1) process is a sequence of random variables $\{X_n\}$ which is formed according to the probabilistic linear model

(2.2) $X_n = U_n Y_{n-D_n} + (1 - U_n) Z_{n-(N+1)}$

for $n = 1,2,\ldots$, where the "autoregressive tail" is

(2.3) $Z_n = V_n Z_{n-A_n} + (1 - V_n) Y_n$

6

for   n = -N-p+1, -N-p+2,....   This process differs

from the DARMA(1,N+1) process defined in Jacobs and Lewis

[1978a] in that the "autoregressive tail" Z_n is now the   pth

order autoregressive process of (1.4).   In Jacobs and Lewis

[1978c], it was shown that the vector-valued Markov chain

$\{(Z_n, Z_{n+1}, \ldots, Z_{n-p+1})$, n = 1,2,...\} has a limiting joint

probability mass function ν with marginal probability mass

function π.   Hence, if $(Z_{-N-p+1}, \ldots, Z_{-N})$ has joint probability

mass function ν , then $\{X_n;$ n = 1,2,...\} is a stationary

process with marginal probability mass function π.

Let   r(k) = Corr(X_n, X_{n+k}) for the stationary process.

Then $\{r(k)\}$ can be shown to satisfy the following system of

equations:

$$(2.4) \qquad r(1) = \beta^2 \sum_{i=0}^{N-1} \delta_i \delta_{i+1} + \beta(1-\beta) r_B(1) + (1-\beta)^2 r_A(1) \ ,$$

$$(2.5) \qquad r(2) = \beta^2 \sum_{i=0}^{N-2} \delta_i \delta_{i+2} + \beta(1-\beta) r_B(2) + (1-\beta)^2 r_A(2) \ ,$$

$$\vdots$$

$$(2.6) \qquad r(N) = \beta^2 \delta_0 \delta_N + \beta(1-\beta) r_B(N) + (1-\beta)^2 r_A(N) \ ,$$

$$(2.7) \qquad r(N+k) = \beta(1-\beta) r_B(N+k) + (1-\beta)^2 r_A(N+k) \ . \qquad k \geq 1.$$

In these equations

$$(2.8) \qquad r_A(k) = \text{Corr}(Z_n, Z_{n+k}) \ , \qquad\qquad k \geq 1,$$

7

which satisfy the following Yule-Walker equations:

$$(2.9) \qquad r_A(1) = \rho\alpha_1 r_A(0) + \rho\alpha_2 r_A(1) + \ldots + \rho\alpha_p r_A(p-1) \ ,$$

$$(2.10) \qquad r_A(2) = \rho\alpha_1 r_A(1) + \rho\alpha_2 r_A(0) + \ldots + \rho\alpha_p r_A(p-2) \ ,$$

$$\cdot \qquad\qquad \cdot \qquad\qquad \cdot \qquad\qquad \cdot$$
$$\cdot \qquad\qquad \cdot \qquad\qquad \cdot \qquad\qquad \cdot$$
$$\cdot \qquad\qquad \cdot \qquad\qquad \cdot \qquad\qquad \cdot$$

$$(2.11) \qquad r_A(p) = \rho\alpha_1 r_A(p-1) + \rho\alpha_2 r_A(p-2) + \ldots + \rho\alpha_p r_A(0) \ ,$$

and for $k \geq 1$,

$$(2.12) \qquad r_A(p+k) = \rho\alpha_1 r_A(p+k-1) + \rho\alpha_2 r_A(p+k-2) + \ldots + \rho\alpha_p r_A(k) \ ,$$

where $r_A(0) = 1$.

In addition

$$r_B(i) = Corr(Z_{n+i-(N+1)}, Y_{n-D_n}) = \sum_{j=0}^{N} Corr(Z_{n+i-(N+1)}, Y_{n-j}) \delta_j$$

is obtained recursively as

$$r_B(0) = 0 \ ;$$

$$r_B(1) = (1-\rho)\delta_N \ ;$$

$$r_B(2) = \rho\alpha_1 r_B(1) + (1-\rho)\delta_{N-1} \ ;$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$r_B(k) = \sum_{i=1}^{k-1} \rho\alpha_i r_B(k-i) + (1-\rho)\ \delta_{N-(k-1)}$$

for $k < min(p,N)$;

$$r_B(k) = \sum_{i=1}^{p} \rho \alpha_i r_B(k-i) + (1-\rho) \, \delta_{N-(k-1)}$$

for $\max(p,N) > k \geq \min(p,N)$;

$$r_B(k) = \sum_{i=1}^{p} \rho \alpha_i r_B(k-i)$$

for $k \geq \max(p,N)$.

To see that the serial correlations for the DARMA$(p,N+1)$ process are all nonnegative let $q(i)$ (respectively $q_A(i)$) be the probability that X_n and X_{n+i} (respectively Z_n and Z_{n+i}) choose the same random variable Y_k , where, because of the backward definition of the autoregression $k \leq n$. Then $q(i)$ (respectively $q_A(i)$) also satisfy equations (2.4) – (2.7) (respectively (2.9) – (2.12)) and since they are nonnegative, so are the serial correlations.

To see this identity, let R_n be the random index of the Y_k, $k \leq n$, that X_n chooses; that is,

$$X_n = Y_{R_n} .$$

Then, since the random variables R_n are independent of the $\{Y_k\}$ random variables,

(2.13) $E\left[X_n X_{n+\ell}\right] = E\left[Y_{R_n} Y_{R_{n+\ell}}\right]$

$$= \sum_{k=1}^{n} E\left[Y_k^2\right] P\{R_n = R_{n+\ell} = k\}$$

$$+ \sum_{k=1}^{n} \sum_{\substack{j=1 \\ j \neq k}}^{n+\ell} E\left[Y_k Y_j\right] P\{R_n = k, R_{n+\ell} = j\}$$

(2.14) $$= E\left[Y_1^2\right] P\{R_n = R_{n+\ell}\}$$

$$+ E\left[Y_1\right]^2 P\{R_n \neq R_{n+\ell}\} \ .$$

Thus

$$\mathrm{Cov}(X_n, X_{n+\ell}) = E\left[Y_1^2\right] P\{R_n = R_{n+\ell}\}$$

$$+ E\left[Y_1\right]^2 \left(P\{R_n \neq R_{n+\ell}\} - 1\right)$$

(2.15) $$= \mathrm{Var}\left[Y_1\right] P\{R_n = R_{n+\ell}\} \ .$$

Therefore

(2.16) $\mathrm{Corr}(X_n, X_{n+\ell}) = P\{R_n = R_{n+\ell}\} = q(\ell)$

as asserted above.

This identity will also be used in the estimation procedure proposed in Section 3 for the serial correlations.

2.2 The NDARMA(p,N) process

In this subsection we will define another related discrete time series with the correlation structure of a mixed moving average autoregressive process. This new process is more reminiscent of the linear ARMA(p,N) process. The key idea that leads to this new model is that a probabilistic mixture of a finite number of random variables each with probability mass function π has probability mass function π even if the random variables are dependent. Thus it is not necessary to define the autoregression via an autoregressive tail, as in the DARMA(p,N+1) process; the autoregression can be made explicit, as in the usual (normal theory) linear processes.

Thus let

$$(2.17) \qquad X_n = V_n X_{n-A_n} + (1-V_n) Y_{n-D_n} \ ,$$

where $\{V_n\}$, $\{A_n\}$, and $\{D_n\}$ are as before. Thus, with probability ρ , X_n is one of the p previous values $X_{n-1}, \ldots, X_{n-p}$ and with probability $(1-\rho)$ it is a mixture of the previous Y_k's , $n - N \leq k \leq n$. Note that if $\rho = 0$, then $\{X_n; n = 1,2,\ldots\}$ is a DMA(N) process as defined in Jacobs and Lewis [1978a]. If $P\{D_n = 0\} = 1$, then $\{X_n\}$ is a DAR(p) process as defined in Jacobs and Lewis [1978c].

Let $\tau = \inf\{i : \delta_i > 0\}$. Note that

$$Z_n = \{(X_n , X_{n-1} , \ldots, X_{n-p+1} , Y_{n-\tau} , \ldots, Y_{n-N}) , n = 1,2,\ldots\}$$

is a Markov Chain with state space $\rm I\!F$ which is equal to the product space of $\rm I\!E$ with itself $p + (N-\tau)$ times. Since

11

$$P\{X_{n+\tau+1} = Y_{n+1} = j \,|\, X_0, \ldots, X_n, Y_0, \ldots, Y_n\} \geq (1-\rho)\,\delta_\tau\,\pi(j),$$

there is a set $J \subseteq \mathbb{F}$ such that $\displaystyle\min_{\substack{i \in \mathbb{F} \\ k \in J}} P\{Z_{n+K} = k \,|\, Z_n = i\} = \gamma > 0$,

where

$K = P + N$ and

$$J \subset \{X_{n+K} = Y_{n+K-\tau},\ X_{n+K-1} = Y_{n+K-1-\tau}, \ldots, X_{n+K-p+1} = Y_{n+K-p+1-\tau}\}\ .$$

Thus the condition of case (b) on page 173 of Doob [1953] is satisfied. The proof on pages 173 and 174 extended to countable state spaces shows that Z_n has a limiting probability mass function ν as $n \to \infty$; further the convergence of the conditional distribution of Z_n to ν as $n \to \infty$ is geometric. The marginal probability mass function of ν is π.

It follows from (2.17) that the serial correlations for the stationary NDARMA(p,N) process satisfy the Yule-Walker equations for the ARMA(p,N) process with restrictions on the range of the coefficients;

(2.19) $r_N(k) \equiv \mathrm{Corr}(X_n, X_{n+k})$

$$= \sum_{i=1}^{p} \rho\alpha_i\, \mathrm{Corr}(X_n, X_{n+k-i})$$

$$+ (1-\rho) \sum_{i=k}^{N} \delta_i\, \mathrm{Corr}(X_n, Y_{n+k-i})$$

for $k \geq 0$. The correlations $r_{NB}(i) = \text{Corr}(X_n , Y_{n-i})$ can be computed recursively as follows. For $i = 0$

(2.20) $$r_{NB}(0) = (1-\rho)\delta_0 ;$$

for $1 \leq i < p$

$$r_{NB}(i) = (1-\rho)\delta_i + \rho\alpha_1 r_{NB}(i-1) + \ldots + \rho\alpha_i r_{nb}(0) ;$$

and for $i > p$

$$r_{NB}(i) = (1-\rho)\delta_i + \sum_{j=1}^{p} \rho\alpha_i r_{NB}(i-j) ,$$

where if $i > N$, then $\delta_i = 0$ by convention.

Hence, if we assume $N < p$

(2.21) $r_N(k) = \rho\alpha_1 r_N(k-1) + \rho\alpha_2 r_N(k-2) + \ldots + \rho\alpha_p r_N(p-k)$

$$+ (1-\rho) \sum_{i=k}^{N} r_{NB}(i-k)$$

for $1 \leq k \leq N$;

(2.22) $r_N(k) = \rho\alpha_1 r_N(k-1) + \rho\alpha_2 r_N(k-2) + \ldots + \rho\alpha_p r_N(k-p)$

for $k > N$.

The serial correlations of the NDARMA(p,N) process are nonnegative since, if $q_N(i)$ is the probability that X_n and X_{n+i} choose the same random variable Y_k , $k \leq n$, then $\{q_N(i)\}$ satisfies equations (2.21) and (2.22). The argument is the same as for the DARMA(p,N+1) case.

2.3 Comparison of Admissible Range of Correlations for the DARMA(1,1) and the NDARMA(1,1).

Let $\{X_n\}$ be a stationary DARMA(1,1) process; that is,

(2.23)
$$X_n = \begin{cases} Y_n & \text{with probability } \beta, \\ Z_{n-1} & \text{with probability } 1-\beta, \end{cases}$$

14

where

(2.24)
$$
Z_n = \begin{cases} Z_{n-1} & \text{with probability } \rho, \\ \\ Y_n & \text{with probability } 1-\rho. \end{cases}
$$

Equations (2.4) - (2.12) for the DARMA(p,N+1) correlations simplify to

(2.25) $r(k) = \text{Corr}(X_n, X_{n+k}) = \rho^{k-1}(1-\beta)[\beta(1-\rho) + (1-\beta)\rho]$.

Similarly let $\{X'_n\}$ be a stationary NDARMA(1,1) process; that is

(2.26) $X'_n = \begin{cases} X'_{n-1} & \text{with probability } \rho, \\ \\ Y_n & \text{with probability } (1-\rho)\delta_0, \\ \\ Y_{n-1} & \text{with probability } (1-\rho)(1-\delta_0) \end{cases}$.

Equations (2.21) and (2.22) simplify to

(2.27) $r_N(k) = \text{Corr}(X'_n, X'_{n+k}) = \rho^{k-1}[\rho + (1-\rho)^2\delta_0(1-\delta_0)]$.

Figure 1 gives graphs of the attainable values of $\{r_N(2), r_N(1)\}$ as the parameter values ρ and δ_0 vary, and $\{r(2), r(1)\}$ as the parameters ρ and β vary. Note that although the set of attainable correlations for the NDARMA(1,1) process is not strictly contained in that for the DARMA(1,1)

15

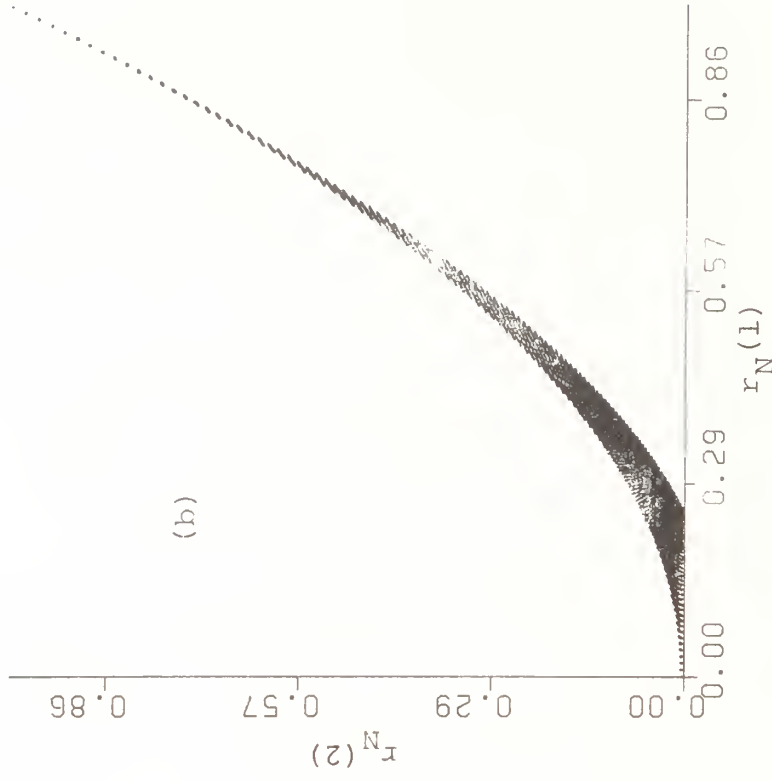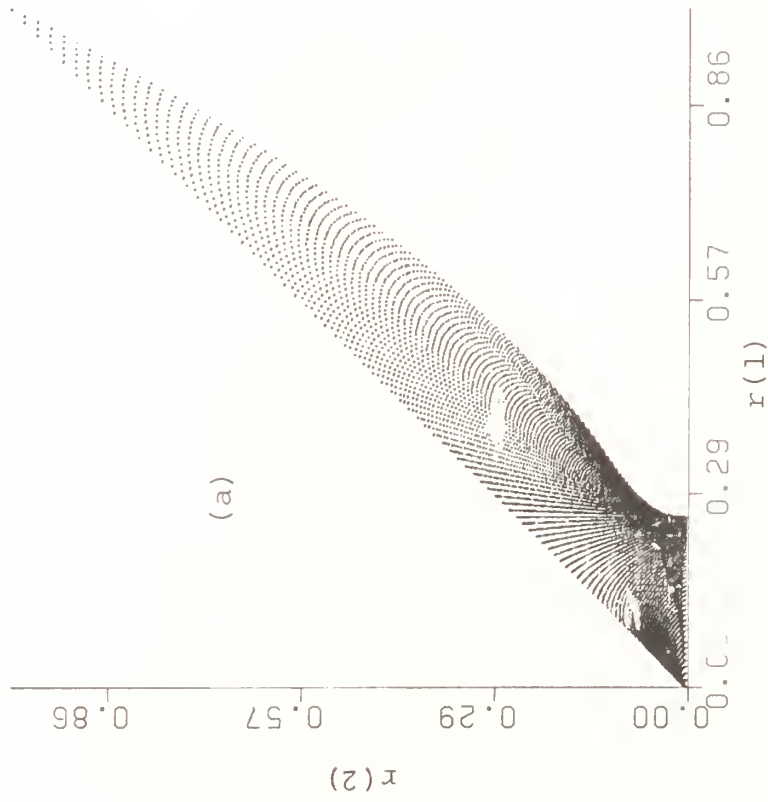Figure 1. Graph (a) of the attainable values of $\{r(2), r(1)\}$ as the parameters ρ and β vary in the DARMA(1,1) process and graph (b) of the attainable values of $\{r_N(2), r_N(1)\}$ as the parameters ρ and δ_0 in the NDARMA(1,1) process vary. Both processes contain the first-order moving average process for which the first serial correlation varies between 0 and 1/4 while the second serial correlation is zero. However, the regions beyond this for the two processes differ considerably. Beyond $r_N(1)$ or $r(1)$ greater than 1/4 the DARMA(1,1) process has a much broader correlation range than the NDARMA(1,1) process. The first lies in a broad band above $r(2) = [r(1)]^2$, the second in a narrow band below $r_N(2) = [r_N(1)]^2$.

process, it is much smaller. Thus, the DARMA(1,1) model appears
to be broader than the NDARMA(1,1) model. The smaller region
of possible correlation pairs for the NDARMA(1,1) model seems to
be a constraint due to the explicit autoregression on X_{n-1}.

2.4 An Estimator for the Serial Correlations of the
DARMA(p,N+1) and NDARMA(p,N) processes.

The usual estimator for the serial correlations of a
stationary real-valued sequence $\{X_1, \ldots, X_m\}$ is

(2.28) $$\hat{r}(\ell) = [s_X^2]^{-1} (m-\ell)^{-1} \sum_{j=1}^{m-\ell} (X_j - \bar{X})(X_{j+\ell} - \bar{X}) ,$$

where

(2.29) $$\bar{X} = m^{-1} \sum_{j=1}^{m} X_j$$

and

(2.30) $$s_X^2 = \frac{1}{m-1} \sum_{j=1}^{m} (X_j - \bar{X})^2 .$$

In this subsection we will suggest another estimator
for the serial correlations of the DARMA and NDARMA processes.

By the remarks at the ends of Sections 2.1 and 2.2,
the ℓth serial correlation, $r(\ell)$, for both the stationary
DARMA(p,N+1) and NDARMA(p,N) processes is equal to the prob-
ability that X_n and $X_{n+\ell}$ choose the same Y_k , $k \leq n$.
Hence, for both processes for $i \neq j$,

17

$$P\{X_n = i, X_{n+\ell} = j\} = P\{Y_{R_n} = i, Y_{R_{n+\ell}} = j\}$$

$$= \sum_{k \le n} \sum_{\substack{r \le n+\ell \\ r \ne k}} P\{R_n = k, R_{n+\ell} = r\} P\{Y_k = i, Y_r = j\}$$

Since the $\{Y_k\}$ random variables are independent, and independent of R_n and $R_{n+\ell}$, we have

$$(2.31) \qquad P\{X_n = i, X_{n+\ell} = j\} = \pi(i)\pi(j)P\{R_n \ne R_{n+\ell}\}$$

$$= \pi(i)\pi(j)[1 - r(\ell)] \ ,$$

by equation (2.16).

Thus, for $j \in \mathbb{E}$

$$(2.32) \quad \lim_{N \to \infty} B_N(m,j) \equiv \lim_{N \to \infty} (N-m)^{-1} \sum_{i \ne j} \sum_{k=1}^{N-m} 1_i(X_k) \ 1_j(X_{k+m})$$

$$= [1-\pi(j)][1-r(m)] \ \pi(j)$$

almost surely where $1_i(x) = 1$ if $x = i$ and 0 otherwise. Hence

$$(2.33) \qquad \tilde{r}(m) = 1 - \sum_{j \in \mathbb{E}} B_N(m,j)[1 - \pi(j)]^{-1}$$

is a strongly consistent estimator for the mth serial correlation for the stationary DARMA and NDARMA processes.

Estimator $\tilde{r}(m)$ is also a strongly consistent estimator for the finite state space models of Lindqvist(1978) and Pegram (1980) since the conditional probabilities (1.2) and (1.6) are of the same form as those for the appropriate DAR(p) process.

In the next section we pursue this estimator for the special case of the first-order autoregressive process DAR(1).

3. ESTIMATION FOR THE DAR(1) PROCESS

Let $\{X_n\}$ be a stationary DAR(1) process with state space $\mathbb{E} = \{0,1,\ldots\}$ and first-order serial correlation ρ, $0 \leq \rho < 1$; that is,

$$(3.1) \qquad X_n = \begin{cases} X_{n-1} & \text{with probability } \rho, \\ \\ Y_n & \text{with probability } 1-\rho \end{cases}$$

for $n = 1,2,\ldots$, while X_0 is a random variable independent of $\{Y_n\}$ but with the same probability mass function.

A simulation was conducted to study the performance of several estimator for ρ for small to moderate series lengths m. The series lengths considered were $m = 20$, 50, and 200. The marginal probability mass functions considered were the Poisson with parameter λ,

$$(3.2) \qquad \pi(k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad k = 0,1,\ldots,$$

and the geometric with parameter p

$$(3.3) \qquad \pi(k) = p^k(1-p) \qquad k = 0,1,\ldots.$$

One type of estimator considered was the single parameter maximum likelihood estimator. For a series of length m let N_{ij} denote the number of times the DAR(1) process goes from i to j, for $i, j \in \mathbb{E}$ and let N_j denote the total

number of times the process is in state j. The log-likelihood function for a DAR(1) series $\{X_1, \ldots, X_m\}$ of length m is

(3.4)
$$L = \sum_{i=0}^{\infty} \sum_{j \neq i} N_{ij} \, \ell n[(1-\rho)\pi(j)]$$

$$+ \sum_{i=0}^{\infty} N_{ii} \, \ell n[1 - (1-\rho)\pi(i)]$$

$$+ \sum_{i=0}^{\infty} 1_i(X_1)\pi(i) \ .$$

Taking the partial derivative of L with respect to $x = 1-\rho$ and setting the derivative equal to zero results, after some simplification, in the following equation for the maximum likelihood estimator $\hat{x} = 1-\hat{\rho}$, if it exists:

(3.5)
$$f(\hat{x}) \equiv 1 - N^{-1} \sum_{i=0}^{\infty} N_{ii} \{1 - \hat{x}[1 - \pi(i)]\}^{-1} = 0 \ .$$

Note that $f(x)$ is monotone decreasing in x and $f(0) \geq 0$. Hence, if there is a solution to (3.5) in $[0,1]$, it will be unique.

The ad-hoc estimator of ρ given at (2.33) was also considered. For the DAR(1) process this estimator is

(3.6)
$$\tilde{\rho} = 1 - \sum_{j=0}^{\infty} \left\{ [N^{-1} \sum_{\substack{i=0 \\ i \neq j}}^{\infty} N_{ij}][1 - \pi(j)]^{-1} \right\} \ .$$

21

The summations in both the maximum likelihood estimators and the estimators of the form (3.6) for the Poisson case (respectively the geometric case) were restricted to be between $a = \max[1, n_-(\mu-7\sigma)]$ (respectively $a = \max(1, n_-(\mu-10\sigma))$ and $b = n_+(\mu+7\sigma)$ (respectively $n_+(\mu+10\sigma)$); here $n_-(y)$ (respective $n_+(y)$) is the largest (respectively smallest) integer less (respectively greater) than y .

Equation (3.5) was solved numerically. In the case $N = 20$, it was not uncommon that $f(x)$ did not have a zero in $[0,1]$. In this case, if $f(1) > 0$, then x was set equal to 1 ; that is, $\hat{\rho} = 0$. If $f(0) = 0$, then x was taken to be 0 ; that is, $\hat{\rho} = 1$.

Other estimators for ρ that were considered included the following:

1. The usual estimate for first-order serial correlation

$$(3.7) \qquad \rho_1 = [S_X^2]^{-1} (m-1)^{-1} \sum_{n=1}^{m-1} (X_n - \bar{X})(X_{n+1} - \bar{X}) ,$$

where $\bar{X}$ and S_X^2 are as in (2.29) and (2.30). If $S_X^2 = 0$, then ρ_1 was set equal to 1 .

2. The maximum likelihood equation (3.5) was solved numerically for ρ for each of the following three values for $\pi(j)$:

a. the known distribution (3.2) or (3.3) with known parameter λ or p was used; ρ_2 denotes this estimator;

b. the known distribution with an estimated parameter was used; that is,

22

(3.8) $\qquad \pi(k) = e^{-\hat{\lambda}} \dfrac{(\hat{\lambda})^k}{k!}$ $\qquad$ where $\qquad \hat{\lambda} = \bar{X}$

or

(3.9) $\qquad \pi(k) = \hat{p}^k (1-\hat{p})$ $\qquad$ where $\qquad \hat{p} = \bar{X}[1 + \bar{X}]^{-1}$;

ρ_3 denotes this estimator;

$\qquad$ c. the nonparametric estimator $\hat{\pi}(j) = N^{-1} N_j$ of $\pi(j)$ was used; ρ_4 denotes this estimator.

$\qquad$ d. ρ_5 is the estimator of ρ resulting from the two-dimensional maximum likelihood estimator where the other parameter is the distribution parameter (λ or p);

$\qquad$ 3. $\qquad$ the estimate ρ_6 is the nonparametric estimate (3.6) using $N^{-1} N_j$ as the estimate for $\pi(j)$;

$\qquad$ 4. $\qquad$ the estimate ρ_7 is the nonparametric estimate (3.6) using the true value of $\pi(j)$;

$\qquad$ Both estimators ρ_6 and ρ_7 can have negative values for small to moderate sample sizes. Hence, we also considered the following estimate.

$\qquad$ 5. $\qquad$ Estimator $\rho_8 = \max(\rho_6 , 0)$.

3.2 The sampling experiment.

$\qquad$ A DAR(1) series of length m was simulated and the estimates for ρ were computed. The computation was repeated for 1000 independent replications and the sample mean, sample variance, and sample root mean square error were computed. Each experiment was then repeated for 20 independent replications,

and the mean of the means, mean of the standard deviations, and mean of the root mean square errors were computed. Tables 1-3 give the means of the root mean square errors for the cases studied. The box plots of row values appearing in the last column of the table are given to help the reader to summarize the performance of the 8 estimators across the 7 cases considered.

All runs were performed on an IBM system 360/67 computer at the Naval Postgraduate School using the LLRANDOM package [Learmonth and Lewis, 1973] which generates numbers according to the scheme given by Lewis, Goodman, and Miller [1969]. Tests of the random number generator are given in Learmonth and Lewis [1974].

Among all the estimates of ρ, the usual first-order serial correlation estimator, ρ_1, performs least well. The maximum likelihood estimator with smallest root mean square error tends to be ρ_2, although by the time $m = 200$ the difference is minor. Of course the value of p or λ would not be known in general, so that this estimator is unrealistic. Maximum likelihood estimators ρ_3 and ρ_5 are about equivalent, indicating that the extra computational complexity of the two-parameter maximum likelihood estimator, ρ_5, is not necessary. The performance of the nonparametric estimator ρ_8 is about the same or sometimes better than that of the maximum likelihood estimator ρ_4, especially if, for small sample size (m=20), the modification $\rho_8 = \max(\rho_6, 0)$ is used. This suggests that the ad-hoc estimator of (3.6) altered to give values in the range [0,1] is almost as good as the maximum likelihood

24

estimator with the same estimate for $\pi(j)$. The ad-hoc estimator is much easier to compute than the maximum likelihood estimator.
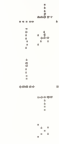
Table 1. Estimated Root Mean Square Error for several
competing estimators of ρ for DAR(1) Series of length
m = 20

Estimator	Poisson λ = 1 ρ=0.05	Poisson λ = 1 ρ=0.60	Poisson λ = 3 ρ=0.50	Poisson λ = 10 ρ=0.75	Geometric p = 0.3925 ρ=0.25	Geometric p = 0.63210 ρ=0.60	Geometric p = 0.63210 ρ=0.95	Box Plots of row Values
1. Sample serial correlation	.220	.277	.267	.290	.256	.295	.273	
2. Maximum likelihood; Known parameter	.110	.145	.133	.105	.173	.135	.056	
3. Maximum likelihood; Parametric est. of $\pi(j)$	.101	.164	.134	.106	.169	.149	.129	
4. Maximum likelihood; Nonparametric est. of $\pi(j)$	.082	.203	.174	.175	.173	.190	.233	
5. Maximum likelihood; Both parameters	.102	.164	.134	.105	.169	.148	.129	
6. Ad hoc estimator; nonparametric est. of $\pi(j)$	.168	.202	.155	.151	.226	.178	.191	
7. Ad hoc estimator; true value of $\pi(j)$	.157	.164	.134	.105	.225	.149	.099	
8. Ad hoc estimator; $\max(\rho_6, 0)$.	.093	.192	.154	.148	.176	.171	.189	

Table 2. Estimated Root Mean Square Error for several
Estimators of ρ for DAR(1) Series of length
m = 50

Estimator	Poisson λ = 1		Poisson λ = 3	Poisson λ 10	Geometric p = 0.3925	Geometric p = 0.63210		Box Plots of row Values
ρ	0.05	0.60	0.50	0.75	0.25	0.60	0.95	
1. Sample serial correlation	.144	.164	.165	.155	.180	.193	.198	
2. Maximum likelihood; Known parameter	.071	.092	.083	.066	.113	.085	.036	
3. Maximum likelihood; Parametric est. of π(j)	.068	.096	.083	.066	.115	.088	.091	
4. Maximum likelihood; Nonparametric est. of π(j)	.061	.106	.094	.081	.119	.098	.154	
5. Maximum likelihood; Both parameters	.068	.095	.084	.066	.115	.087	.090	
6. Ad hoc estimator; nonparametric est. of π(j)	.101	.101	.088	.071	.127	.092	.132	
7. Ad hoc estimator; true value of π(j)	.095	.095	.084	.066	.130	.087	.071	
8. Ad hoc estimator; max(ρ_6,0) .	.068	.101	.088	.072	.121	.092	.131	

Table 3. Estimated Root Mean Square Error for several Estimators of ρ for DAR(1) Series of Length m = 200

Estimator	Poisson λ = 1		Poisson λ = 3	Poisson λ = 10	Geometric p = .3925	Geometric p = .63210		Box Plots of row Values
ρ	0.05	0.60	0.50	0.75	0.25	0.60	0.95	
1. Sample serial correlation	.075	.082	.082	.068	.103	.103	.064	
2. Maximum likelihood; Known parameter	.040	.046	.042	.033	.056	.042	.020	
3. Maximum likelihood; Parametric est. of π(j)	.039	.047	.042	.033	.057	.043	.023	
4. Maximum likelihood; Nonparametric est. of π(j)	.038	.048	.043	.035	.058	.045	.029	
5. Maximum likelihood; Both parameters	.039	.047	.042	.033	.060	.043	.023	
6. Ad hoc estimator; nonparametric est. of π(j)	.049	.047	.042	.034	.061	.044	.028	
7. Ad hoc estimator; true value of π(j)	.049	.047	.042	.033	.064	.043	.023	
8. Ad hoc estimator; max(ρ_6,0) .	.042	.047	.042	.034	.061	.044	.028	

4. EXTENSIONS

Since the process $\{X_n\}$ is obtained as a probabilistic mixture of the $\{Y_n\}$, the DARMA and NDARMA process may be defined using any sequence of independent identically distributed random variables $\{Y_n\}$. One implication is that DARMA and NDARMA processes may have a continuous marginal distribution. However, even if the distribution of Y_n is continuous, a realization of the sequence $\{X_n\}$ will, in general, contain many runs of a single value. This seems to be the major drawback to using DARMA and NDARMA processes to obtain a sequence of dependent random variables with a specified continuous distribution and correlation structure. However, the process with continuous marginals may be useful in simulation studies.

Multivariate DARMA and NDARMA processes may be obtained by using a sequence of multivariate Y_n's . To illustrate this we generate DARMA and NDARMA-like processes having negative correlations. These can be derived from bivariate processes as follows.

Let $\{(Y_n(1), Y_n(-1))\}$ be a sequence of independent bivariate random variables with state space $\mathbb{E} = \{0, \pm 1, \ldots\}$, marginal probability mass function π , and correlation $r = \mathrm{Corr}(Y_n(1), Y_n(-1))$ which will be negative in general. One way to generate such a sequence is to note that a random variable $Y_n(1)$ with probability mass function π can be simulated from a uniform $[0,1]$ random variable by defining

$$(4.1) \qquad Y_n(1) = j \qquad \text{if} \qquad \sum_{i=-\infty}^{j-1} \pi(i) < U \leq \sum_{i=-\infty}^{j} \pi(i) \ .$$

If $Y_n(-1)$ is generated by

$$(4.2) \qquad Y_n(-1) = j \qquad \text{if} \qquad \sum_{i=-\infty}^{j-1} \pi(i) < 1 - U \leq \sum_{i=-\infty}^{j} \pi(i) \ ,$$

then $(Y_n(1), Y_n(-1))$ is called an antithetic pair. If π is symmetric about zero, then $Y_n(1) = - Y_n(-1)$, and $r = \mathrm{Corr}(Y_n(1), Y_n(-1)) = -1$.

A bivariate DARMA $(p, N+1)$ process $\{(X_n(1), X_n(-1))\}$ is defined as follows. Let $\{a_0, \ldots, a_p\}$ and $\{b_0, \ldots, b_N\}$ be fixed sequences of numbers that are either -1 or 1 . Let

$$(4.3) \qquad X_n(1) = U_n Y_{n-D_n}(b_{D_n}) + (1-U_n) Z_{n-(N+1)}(a_0)$$

$$(4.4) \qquad X_n(-1) = U_n Y_{n-D_n}(-b_{D_n}) + (1-U_n) Z_{n-(N+1)}(-a_0)$$

for $n = 1, 2, \ldots$,

where

$$(4.5) \qquad Z_n(a_0) = V_n Z_{n-A_n}(a_{A_n}) + (1-V_n) Y_n(a_0)$$

$$(4.6) \qquad Z_n(-a_0) = V_n Z_{n-A_n}(-a_{A_n}) + (1-V_n) Y_n(-a_0)$$

for $n = -N, -N+1, \ldots$ where $\{A_n\}$ and $\{D_n\}$ are as in Section 2. The random variable $X_n(-1)$ is called the dual of $X_n(1)$

or the antithetic when (4.1) and (4.2) hold for the $\{Y_n\}$ pair.
Note that if (4.1) and (4.2) hold and π is symmetric about
zero, then $Z_n(-1) = -Z_n(1)$ and $X_n(-1) = -X_n(1)$.

A bivariate NDARMA(p,q) process $(X'_n(1), X'_n(-1))$ can
be defined similarly:

(4.7) $\qquad X'_n(1) = V_n X'_{n-A_n}(a_{A_n}) + (1-V_n) Y_{n-D_n}(b_{D_n})$,

(4.8) $\qquad X'_n(-1) = V_n X'_{n-A_n}(-a_{A_n}) + (1-V_n) Y_{n-D_n}(-b_{D_n})$.

The stationary bivariate DARMA and NDARMA processes
will have marginal probability mass function π . A process
having possibly negative correlations can be obtained by consider
the marginal processes $\{X_n(1)\}$, $\{X_n(-1)\}$, $\{X'_n(1)\}$, $\{X'_n(-1)\}$.
Details will be given elsewhere.

.

5.   ACKNOWLEDGEMENTS

REFERENCES

BHAT, U. N. Elements of Applied Stochastic Processes. J. Wiley and Sons, New York (1972).

BUISHAND, T. A. The binary DARMA(1,1) process as a model for wet-dry sequences. Technical Note 78-01, Dept. of Mathematics, Statistics Division, Agricultural University Wageningen, The Netherlands (1978).

DOOB, J. L. Stochastic Processes. John Wiley and Sons, New York (1953).

JACOBS, P. A. and LEWIS, P. A. W. Discrete time series generated by mixtures I: correlational and runs properties. J. R. Statist. Soc. B 40 (1978a) 94-105.

_____. Discrete time series generated by mixtures II: asymptotic properties. J. R. Statist. Soc. 40 (1978b) 222-228.

_____. Discrete time series generated by mixtures III: autoregressive processes (DAR(p)). Naval Postgraduate School Technical Report NPS55-78-022, 1978c.

LEARMONTH, G. P. and LEWIS, P. A. W. Naval Postgraduate School random number generator package. LLRANDOM. Naval Postgraduate School Technical Report NPS55Lw 73061A, (1973).

LEARMONTH, G. P. and LEWIS, P. A. W. Statistical tests of some widely used and recently proposed uniform number generators. Proc. Seventh Conference on Computer Science and Statistics, Western Periodicals, North Hollywood, Calif. 1974.

LEWIS, P. A. W., GOODMAN, A. S. and MILLER, J. M. A pseudo random number generator for the IBM System/360. IBM Systems Journal 8, (1969) 136-146.

LINDQVIST, B. A model for multinomial trials with dependence. Statistical Research Report, Institute of Mathematics, University of Oslo, 1978.

PEGRAM, G. G. S. An autoregressive model for multilag Markov chains. J. Appl. Prob. 17 (1980) 350-362.

DISTRIBUTION LIST

NO. OF COPIES

Library, Code 0142   4
Naval Postgraduate School
Monterey, CA 93940

Dean of Research   1
Code 012A
Naval Postgraduate School
Monterey, CA 93940

Library, Code 55   2
Naval Postgraduate School
Monterey, CA 93940

Professor P. A. W. Lewis   250
Code 55Lw
Naval Postgraduate School
Monterey, CA 93940

Defense Technical Information Center
Cameron Station
Alexandria,Virginia 22314   2